

# Generalized cofactors and decomposition of Boolean satisfiability problems

Madhav Desai                      Virendra Sule  
 madhav@ee.iitb.ac.in            vrs@ee.iitb.ac.in

Department of Electrical Engineering  
 Indian Institute of Technology Bombay, India

December 9, 2014

## Abstract

We propose an approach for decomposing Boolean satisfiability problems while extending recent results of [12] on solving Boolean systems of equations. Developments in [12] were aimed at the expansion of functions  $f$  in orthonormal (ON) sets of base functions as a generalization of the Boole-Shannon expansion and the derivation of the consistency condition for the equation  $f = 0$  in terms of the expansion co-efficients. In this paper, we further extend the Boole-Shannon expansion over an arbitrary set of base functions and derive the consistency condition for  $f = 1$ . The generalization of the Boole-Shannon formula presented in this paper is in terms of *cofactors* as co-efficients with respect to a set of CNFs called a *base* which appear in a given Boolean CNF formula itself. This approach results in a novel parallel algorithm for decomposition of a CNF formula and computation of all satisfying assignments when they exist by using the given data set of CNFs itself as the base.

## 1 Introduction

The solution of systems of Boolean equations is one of the fundamental problems of Computer Science several of whose specialized formulations such as CNF satisfiability (SAT) have been under intense study since over the last fifty years [4, 5]. This is due to the vast applications of this problem in engineering, such as in the verification of software and computer hardware design, logic and artificial intelligence, analogy with several Computer Science problems [6, 7] and in recent times to Biology [8]. The 3-SAT problem is also well known as a model NP complete problem. In recent times there has been much interest in development of parallel solvers for the CNF SAT problem [9]. For a scalable parallel solution it is important to have a strategy of decomposition of the CNF formula. A purpose of this paper is to arrive at such a strategy from the generalization of previous results on the well known Boole-Shannon expansion formula for Boolean functions.

Essentially, in this paper we attempt to resolve the CNF-SAT problem and propose an algorithm which has its origin in a generalization of the well known Boole-Shannon formula which has been a basis of Boolean computation [4]. In [11] a consistency condition for an equation  $f(X) = 0$  was first proved using an extension of this formula in terms of an expansion

of  $f$  using a special basis set of what are called as *orthonormal* (ON) functions. While this condition held true over general Boolean algebras the Boolean function  $f$  had to be in the span of the ON set. In [12] this condition was then further extended to cases when  $f$  is not necessarily in the span of the ON set but then worked only for Boolean functions over the algebra  $B_0 = \{0, 1\}$  which is of interest in applications.

In the present paper we first establish that the general Boole-Shannon expansion of functions relative to an ON set can in fact be proved even without orthonormality of the base set but with certain restrictions on  $f$  relative to the chosen base set. This generalization actually stems from an observation that the expansion co-efficients relative to an ON set are just the *cofactors* and satisfy the same algebraic constraints even for general base sets. The resulting consistency condition for the equation  $f(X) = 1$  in terms of expansion co-efficients of  $f$  with respect to an arbitrary base set then generalizes the results of [12]. We then propose an algorithm for CNF SAT using this condition and its consequent decomposition of the CNF formula. This decomposition gives rise to possibilities of efficient parallel solution to the SAT problem.

For the sake of comprehensibility we shall avoid introducing the background on Boolean algebra and functions and shall direct the reader to basic references [3, 2] and the previous paper [12] which initiates the present line of investigation.

## 2 Expansion formulas

We present a brief background of the Boole-Shannon expansion and the generalized orthonormal expansion. We shall always be concerned with Boolean functions  $f(X) : B_0^n \rightarrow B_0$  of  $n$  variables denoted by  $X$ . The set of Boolean functions shall be denoted by  $B_0(X)$ . For an index  $i$  the well known Boole-Shannon formula for  $f$  is the following, which was in essence even known to Boole [1] while Shannon re-established its role in the 20th century,

$$f(X) = x_i f(x_i = 1) + x'_i f(x_i = 0)$$

This formula has a generalization in terms of expansion with respect to ON functions. A set of functions  $\{\phi_1(X), \dots, \phi_m(X)\}$  is said to be orthogonal if, for  $i \neq j$ ,  $\phi_i \phi_j = 0$  the zero function. The set is orthonormal (ON), when in addition

$$\sum_{i=1}^m \phi_i(X) = 1$$

It is easily proved and shown in [2] that any Boolean function  $f$  has an ON expansion

$$f(X) = \sum_{i=1}^m \alpha_i(X) \phi_i(X) \tag{1}$$

where the expansion co-efficient functions  $\alpha_i(X)$  satisfy the interval

$$\alpha_i \in [f\phi_i, f + \phi'_i]$$

The central theme of the previous paper [12] was to determine a consistency condition for the equation  $f(X) = 0$  in terms of the co-efficients  $\alpha_i$ . This paper shall report further generalizations and results.

## 2.1 Expansion in ON terms

The ON expansion shown in Eq. (1) has the well known special case in which the ON functions are ON terms (products). If  $T = \{t_1, \dots, t_m\}$  are ON terms in variables  $X$  then any Boolean function  $f(X)$  has the dual expansions [2, Theorem 3.15.1]

$$\begin{aligned} f(X) &= \sum_{i=1}^m (f/t_i)(X) t_i(X) \\ f(X) &= \prod_{i=1}^m [(f/t_i)(X) + t_i(X)'] \end{aligned}$$

where  $(f/t)(X)$  for any term  $t(X)$  in  $n$  variables  $X$  is called the *quotient* which is defined in the literature [2, 10] as the function  $f(X)$  along with the constraints on variables defined by the equation

$$t(X) = 1$$

Clearly for a term  $t$  the constraint  $t = 1$  results in *partial assignments* of variables in  $X$  which we shall denote as  $q(t)$ . For instance if  $X = \{x_1, x_2, x_3\}$  then for  $t = x_1 x'_3$ ,  $q(t) = \{x_1 = 1, x_3 = 0\}$ . Hence  $f/t = f(1, x_2, 0)$  is a function of  $x_2$ . The definition of quotient thus does not extend to the case when  $t$  is replaced by a sum of terms.

## 2.2 Generalization to cofactors

An identical concept to the quotient  $f/t$  talked about in literature [10] is the concept of *cofactor* of a function  $f$  relative to a term  $t$  denoted as  $f_t$ . However it is defined again as the function  $f$  considered along with the constraint equation  $t = 1$ . We shall thus extend this definition for the general case when the term  $t$  is another function.

**Definition 1.** Let  $f(X), g(X) \neq 0$  be Boolean functions in  $n$  variables  $X$  in  $B_0(X)$ . Then a set of *cofactors* of  $f$  relative to  $g$  denoted  $\Xi(f, g)$  is the set of all Boolean functions whose restriction on the set  $\text{supp}(g)$ <sup>1</sup> is identical to that of the restriction of  $f$  on  $\text{supp}(g)$ .

The cofactors are thus characterized as follows

**Lemma 1.** The set of cofactors  $\Xi(f, g)$  is the set of functions  $\alpha(X)$  which satisfy the equation

$$\alpha(X)g(X) = f(X)g(X) \tag{2}$$

Hence  $\Xi(f, g)$  is given by the interval

$$[f(X)g(X), f(X) + g(X)']$$

and in parametrized form

$$\Xi(f, g) = \{fg + pg', p \in B_0(X)\}$$

*Proof:* If any function  $\alpha(X)$  is in  $\Xi(f, g)$  then  $\alpha g = fg$ . Conversely if  $\alpha$  is a solution of this equation in  $B_0(X)$  then  $\alpha(x) = f(x)$  for  $x$  in  $\text{supp}(g)$ . The interval shown is the set of all solutions  $\alpha$  of this equation. Hence this interval is the set of all cofactors. Every element of this interval is of the form  $fg + pg'$  for an arbitrary function  $p$ .  $\square$

The set  $\Xi(f, g)$  has following properties w.r.t. Boolean operations.

---

<sup>1</sup> $\text{supp } g = \{x \in B_0^n | g(x) = 1\}$

**Proposition 1.** Let  $\alpha, \beta$  belong to  $\Xi(f, g)$  then

1.  $\alpha + \beta$  belongs to  $\Xi(f, g)$ .
2.  $(\alpha)(\beta)$  belongs to  $\Xi(f, g)$ .
3.  $\alpha'$  belongs to  $\Xi(f', g)$  given by the interval  $[f'g, f' + g']$  and has the general form

$$\alpha' = f'g + pg'$$

for an arbitrary function  $p$ .

Note that the preceding proposition implies that  $|\Xi(f, g)| = |\Xi(f', g)|$ .

### 2.3 Representation in cofactors: beyond orthonormal expansion

Now consider the ON set  $g, g'$  and let

$$f = (\alpha)g + (\beta)g'$$

be an ON expansion of  $f$  in this set. Then it is well known that  $\alpha, \beta$  satisfy the equations

$$\begin{aligned} (\alpha)g &= fg \\ (\beta)g' &= fg' \end{aligned}$$

Hence  $\alpha$  belongs to  $\Xi(f, g)$  and  $\beta$  belongs to  $\Xi(f, g')$ . Hence these are just cofactors with respect to an ON set  $\{g, g'\}$ . From this observation we can explore generalization of the expansion formulas as follows.

#### 2.3.1 Boole-Shannon expansion

The well known Boole-Shannon expansion formula for Boolean functions  $f : B_0^n \rightarrow B_0$  in  $n$ -variables  $X$  is

$$f = x_i f(x_i = 1) + x'_i f(x_i = 0)$$

This is generalized by the ON expansion (1) above. However in terms of the cofactors of  $f$  with respect to  $g, g'$ , a further generalization of Boole-Shannon expansion formula is as established above

**Proposition 2.**

$$f = (\alpha)g + (\beta)g' \tag{3}$$

for any  $\alpha \in \Xi(f, g)$  and  $\beta \in \Xi(f, g')$ .

The above expression replaces the ON set  $\{x, x'\}$  in one of the variables to the ON set  $\{g, g'\}$ . The concept of cofactors however allows a further generalization even when the base set is not ON. This is developed in the next subsection.

### 2.3.2 Beyond orthonormality

Consider the set  $\{g, h\}$  of two Boolean functions. Let  $f$  be a given Boolean function whose cofactors w.r.t. this set are  $\Xi(f, g)$  and  $\Xi(f, h)$ . In analogy with the ON expansion formula for  $f$  in terms of  $g, g'$ , we ask, what should be conditions on  $g, h$  such that the function

$$F = (\alpha)g + (\beta)h$$

is equal to  $f$  for any choice of  $\alpha \in \Xi(f, g)$  and  $\beta \in \Xi(f, h)$ ?

Since for any  $x$  in  $B_0^n$  and the fact that  $\alpha, \beta$  are cofactors as above

$$F(x) = g(x)f(x) + h(x)f(x) = (g(x) + h(x))f(x)$$

Hence it follows that  $F = f$  iff

$$f \leq g + h$$

we can thus look for a general statement as in

**Theorem 1.** Let  $G = \{g_1, \dots, g_m\}$ ,  $g_i \neq 0$  for any  $i$ , be a set of Boolean functions whose sum is

$$g = \sum_{i=1}^m g_i$$

Then for any Boolean function  $f \leq g$

$$f = \sum_{i=1}^m \alpha_i g_i \tag{4}$$

for any  $\alpha_i \in \Xi(f, g_i)$ .

*Proof:* Let

$$F = \sum_{i=1}^m \alpha_i g_i$$

where  $\alpha_i$  in  $\Xi(f, g_i)$  be chosen arbitrarily. Then for any  $x$  in  $B_0^n$ ,

$$F(x) = \sum_i f(x)g_i(x) = f(x)\left(\sum_i g_i(x)\right) = f(x)g(x)$$

Hence  $F = fg$ . Since  $f \leq g$  it follows that  $F = f$ .  $\square$

We call this last expansion (4) as *generalized cofactor expansion* of  $f$  relative to  $G$  and it is quite apparent that this generalizes the ON expansion formula (1). This expansion is then a vast generalization of the Boole-Shannon formula going beyond orthonormality.

**Example 1.** Consider  $f = x'_1x_2 + x_2x_3 + x_1x'_3$  and  $C = (x'_1 + x_3)$  then

$$fC = x'_1x_2 + x_2x_3 = x_2C = x'_1x_2 + x_2x_3$$

which is the minimal co-factor in  $\Xi(f, C)$  while

$$f + C' = f$$

and hence the maximal cofactor is  $f$  itself. A general function in  $\Xi(f, C)$  is

$$fC + pC' = x_2(x'_1 + x_3) + (p)(x_1x'_3)$$

where  $p$  is an arbitrary Boolean function in the three variables.

## 2.4 Algebra in terms of cofactors

As in the case of the ON expansion of functions, various computations and operations in the Boolean algebra of functions can be expressed in terms of co-factors since the equation defining the co-factors is the same as that of the equation determining the expansion co-efficients of an ON expansion. For the purpose of completeness these identities are presented below.

**Proposition 3.** Let  $G = \{g_1, \dots, g_m\}$  be a base set as in the above theorem and  $f, h$  be Boolean functions with cofactors relative to  $G$  denoted as  $\alpha_i \in \Xi(f, g_i)$  and  $\beta_i \in \Xi(h, g_i)$ . Then

$$\begin{aligned}\alpha_i + \beta_i &\in \Xi(f + h, g_i) \\ (\alpha)(\beta) &\in \Xi(fh, g_i) \\ \alpha'_i &\in \Xi(f', g_i) \\ \alpha_i \oplus \beta_i &\in \Xi(f \oplus h, g_i)\end{aligned}$$

*Proof:* The proofs follow easily from the characterization of  $\Xi(f, g_i)$  and  $\Xi(h, g_i)$ .  $\square$

In fact the above relations prove that the following identities hold similar to the ON expansion formulas in [2, section 3.14.1]

$$f(X) + h(X) = \sum_{g_i \in G} (\alpha_i(X) + \beta_i(X))g_i(X) \quad (5)$$

$$f(X)h(X) = \sum_{g_i \in G} (\alpha_i(X)\beta_i(X))g_i(X) \quad (6)$$

$$f(X)' = \sum_{g_i \in G} \alpha_i(X)'g_i(X) \quad (7)$$

$$f(X) \oplus g(X) = \sum_{g_i \in G} (\alpha_i(X) \oplus \beta_i(X))g_i(X) \quad (8)$$

Composition of Boolean functions can also be expressed in terms of above expansion formulas as follows. The simple proof is omitted. Consider the base set  $G$  as above and let  $h_i(X)$  for  $i = 1, \dots, n$ ,  $f(X)$  be Boolean functions in  $n$ -variables and

$$h_i(X) = \sum_j \beta_{ij}(X)g_j(X)$$

be expansions of  $h_i(X)$  in the set  $G$  as above. Then

$$f(h_1(X), \dots, h_n(X)) = \sum_j \alpha_j(\beta_{1j}(X), \dots, \beta_{nj}(X))\phi_j(X)$$

## 3 Consistency conditions from the generalized cofactor expansion

As shown in [12] the ON representation despite its non-unique coefficients is useful for solving Boolean equations and from a theoretical standpoint even for deriving consistency conditions on  $f = 0$  when an arbitrary ON expansion of  $f$  is known. In fact it is shown in [12] that the

problem of satisfiability of systems of Boolean equations and that of determining all solutions of such systems can be solved by ON expansion of component functions without the need for forming a single equation from the system.

In view of the above generalization of the ON expansion formula in terms of co-factors it is instructive to explore as far as feasible, the extension of the program followed in [12] in terms of cofactors associated with systems of Boolean equations. We begin this development with a preliminary result.

**Theorem 2.** Let  $G = \{g_1, \dots, g_m\}$  be a set of non-zero Boolean functions

$$g = \sum_{i=1}^m g_i$$

and  $f \leq g$ . Then  $f = 1$  is consistent iff there exists an index  $i \in [1, m]$  such that the system

$$\alpha_i = 1, g_i = 1$$

is consistent for any cofactor  $\alpha_i \in \Xi(f, g_i)$ .

*Proof:* If  $f(x) = 1$  for some  $x$  then it follows from the expansion (4) that for some  $i$ ,  $\alpha_i(x)g_i(x) = 1$ . Hence  $x$  is in  $\text{supp } g$  and  $\alpha_i(x) = 1$  and thus the system is consistent and this is independent of the choice of  $\alpha_i$  in  $\Xi(f, g_i)$ . Conversely if  $f = 1$  is not consistent then  $f(x) = 0$  for all  $x$  in  $B_0^n$ . Hence for any  $x$ , any  $i$  and a choice of cofactor  $\alpha_i$ ,  $\alpha_i(x)g_i(x) = 0$ . This shows that the system  $\alpha_i = 1, g_i = 1$  is not consistent.  $\square$

In terms of an ON base set we can obtain the following result. Let  $\Phi = \{\phi_1, \dots, \phi_m\}$  be an ON set of non-zero functions. Then for any Boolean function  $f$  we have the cofactor based expansion of the form (4)

$$f = \sum_{i=1}^m \alpha_i \phi_i$$

where  $\alpha_i \in \Xi(f, \phi_i)$ .

**Corollary 1.** The equation  $f(X) = 1$  is consistent iff there is exactly one index  $i$  such that the system

$$\alpha_i(X) = 1, \phi_i(X) = 1$$

is consistent for an arbitrary choice of co-factor in the above expansion.

*Proof:* An ON expansion of  $f$  in the base set  $\Phi$  based on cofactors, of the form

$$f = \sum_{i=1}^m \alpha_i \phi_i$$

exists as sum of functions in  $\Phi$  is 1. Since  $\Phi$  is ON and the functions are non-zero, if the equation is consistent and  $f(x) = 1$  for some  $x$ . Since there is exactly one index  $i$  for which the function  $\phi_i(x) = 1$  for any  $x$  while all others vanish at this  $x$  it follows that the given system is consistent for this  $i$  irrespective of the co-factor  $\alpha_i$ . This proves necessity.

Conversely if the given system is consistent for some  $i$  and an  $x$  is a solution, then as  $\phi_j$  are orthogonal to  $\phi_i$  we get  $\phi_i(x) = 1$  and  $\phi_j(x) = 0$  for  $j \neq i$ . Since  $\alpha_i = f\phi_i + p\phi'_i$ , this implies  $f(x) = 1$  for this  $x$  which proves sufficiency.  $\square$

## 4 Decomposition of CNF formulas

Consistency results in the form shown above should be useful for the solution of a variety of problems of Boolean satisfiability. In particular we explore decomposing a CNF formula for satisfiability. Let

$$f = \prod_{i=1}^m C_i$$

be a CNF formula expressed as a product of CNFs. Then the condition

$$f \leq \sum_{i=1}^m C_i$$

holds. Hence we have a representation of  $f$  as in (4)

$$f = \sum_{i=1}^m \alpha_i C_i$$

where  $\alpha_i$  belong to  $\Xi(f, C_i)$ . However we show next that the actual computation of or choice of  $\alpha_i$  for each  $C_i$  can be avoided while theorem 2 can be still utilized for exploring satisfiability.

### 4.1 Partial assignments and their substitution

To explain the application of theorem 2 to the CNF SAT problem we need to introduce the notion of *partial assignment* defined by a clause and notation for substitution of partial assignments in functions. Consider a clause,

$$C = x_{i_1}^{a_1} + \dots + x_{i_k}^{a_k}$$

with  $a_i = 0, 1$  defined by the rule  $x^1 = x$ ,  $x^0 = x'$ . Then  $C = 1$  for any of the assignments of variables in the set  $s(C) = \{x_{i_j} = a_j\}$ . This set is called SAT set of  $C$ . Hence the set denoted  $q(C)$  called the set of partial assignments defined by  $C$ , which consists of all non empty subsets of  $s(C)$ , is the set of assignments of arguments in  $C$  which lead to evaluate  $C = 1$ . For example the CNF

$$C = x'_1 + x_2 + x'_3$$

has the SAT set  $s(C) = \{x_1 = 0, x_2 = 1, x_3 = 0\}$ . Hence the set  $q(C)$  of all partial assignments consists of subsets denoted by columns of the table

$x_1$	0			0	0		0
$x_2$		1		1		1	1
$x_3$			1		1	1	1

For a CNF containing  $k$  terms therefore there are

$$\sum_{r=1}^k \binom{k}{r}$$

elements in  $q(C)$ .

If  $C$  is one of the clauses in a product formula  $f = C_1 C_2 \dots C_m$  then for any fixed  $i$ , substitution of each partial assignment from  $q(C_i)$  results into reduced CNF formulas.



## 4.2 Satisfiability by decomposition

By a reformulation of theorem 2 we get a decomposition of a CNF formula. Let

$$f = \prod_{i=1}^m C_i$$

be a CNF formula.

**Corollary 2.**  $f$  is SAT (i.e.  $f = 1$  is consistent) iff there exists an index  $i$  and a partial assignment  $q$  in  $q(C_i)$  such that the CNF formula

$$f(q)$$

after partial assignment  $q$  is substituted in  $f$  is SAT.

*Proof:* Let  $\alpha_i$  be a cofactor in  $\Xi(f, C_i)$ . Then by theorem 2,  $f$  is SAT iff for some  $i$  there is  $q$  in  $q(C_i)$  such that  $\alpha_i(q) = 1$  is satisfiable. Since partial assignment of variables at this  $q$  gives  $\alpha_i(q, z) = f(q, z)$  for other unassigned variables  $z$ , it follows that the condition is equivalent to satisfiability of  $f(q, z)$ .  $\square$

For each partial assignment  $q$  there is the CNF formula  $f(q)$  whose satisfiability needs to be tested to determine satisfiability of  $f$ . However  $f(q)$  is a reduced formula compared to  $f$  and for different  $q$  these are independent SAT problems. This gives rise to a decomposition of the original SAT problem into reduced problems which can be solved in parallel. An algorithm based on this decomposition is developed in the next section. We conclude with an illustrative example.

**Example 2.** Consider the CNF formula

$$f = C_1 C_2 C_3 C_4 = (x' + y + w)(y' + z + w')(x + z + w')(x + z' + w')$$

The SAT sets of the CNF factors are as follows

$$\begin{aligned} s(C_1) &= \{x = 0, y = 1, w = 1\} \\ s(C_2) &= \{y = 0, z = 1, w = 0\} \\ s(C_3) &= \{x = 1, z = 1, w = 0\} \\ s(C_4) &= \{x = 1, z = 0, w = 0\} \end{aligned}$$

The set  $q(C_1)$  is thus shown by the indexed columns of the table

	1	2	3	4	5	6	7
$x$	0			0	0		0
$y$		1		1		1	1
$w$			1		1	1	1

Reduced CNFs at various partial assignments in  $q(C_1)$  are shown below indexed by the columns in the above table

- 1  $(y' + z + w)(z + w')(z' + w')$
- 2  $(z + w')(x + z + w')(x + z' + w')$
- 3  $(y' + z)(x + z)(x + z')$
- 4  $(z + w')(z' + w')$
- 5  $(y' + z)(z)(z')$
- 6  $(z)(x + z)(x + z')$
- 7  $(z)(z')$

Similarly reduced CNFs at  $q(C_i)$  for  $i = 2, 3, 4$  can be computed.

### 4.3 All solutions of a CNF formula

When a CNF formula denoted  $F(C)$  which is the product of all CNFs in a set  $C$  of CNFs is satisfiable, the set  $L(C)$  of all solutions of  $F = 1$  (also called the SAT solutions of  $C$ ) is the intersection of all partial assignments satisfying CNFs in  $C$ . Representing such a set of solutions efficiently is a problem of developing a suitable data structure. We consider next the problem of decomposing a formula  $F$  in terms of smaller formulas using the decomposition facilitated by the above theory. The problem of returning all solutions of a formula is yet another problem which shall be addressed elsewhere.

### 4.4 Algorithm for CNF SAT

The idea of the algorithm to be described now follows from the example of the last section. The algorithm identifies parallel computations due to the inherent decomposition available and is built using following functions. We consider  $C$  a set of given CNFs,  $F(C)$  the formula which is the product of all CNFs in  $C$ . The algorithm utilizes following functions

1. The function **Decompose()** which decomposes a given data set  $C$  of CNFs in terms of sets of smaller size from which all solutions can be computed by solving all solutions for each smaller set.
2. A function **Varlist()** extracts and indexes the variable list inside a set  $C$ .
3.  $n_0$  the smallest number of variables in  $C$  at which further decomposition is not utilized for computation of solutions.
4. Function **Partial-assignments()** returns all partial assignments satisfying a set  $C$ .

**function** DECOMPOSE( $C, X, n_0$ )

**Data**  $C$  Set of CNFs,  $X$  variable list in  $C$ ,  $n_0$  number.

**Output** CNF sets with number of variables  $\leq n_0$  whose union of solutions is equal to the set of all SAT solutions of  $C$ .

**loop%** The main Loop

**while**  $|X| > n_0$  **do**

    Choose a subset  $X_1 \subset X$  of variables,  $|X_1| \leq n_0$ .

    Partition  $C = C_1 \sqcup C_2 \sqcup C_3$ ,  $X = X_1 \sqcup X_2$

    % $C_1$  set of clauses containing only  $X_1$  variables.

    % $C_2$  set of clauses containing elements of both  $X_1$  and  $X_2$ .

    % $C_3$  set of clauses containing only  $X_2$  variables.

    Compute all partial assignments  $L = \mathbf{Partial-assignment}(C_1)$ .

**if**  $L = \emptyset$  **then**

        Unsatisfiable formula.

        Exit with output UN-SAT

**else**

        Gather all partial assignments at a memory element.

        Identify indices of variables assigned.

```

    end if
    for  $q_i \in L$  do
        Compute  $C_2(q_i)$ 
        Send  $C_2(q_i)$  to independent processing elements  $PE(i)$ .
        Send  $C_3$  to an independent processing element  $PE(|L| + 1)$ .
    end for
end while
Return all satisfying assignments of  $C$ .
end loop
Exit the algorithm.
for each  $PE(i)$  do
     $C \leftarrow C_2(q_i)$ 
     $X \leftarrow \mathbf{Varlist}(C)$ 
    Repeat the main loop at the  $PE(i)$ 
end for
for the processing element  $PE(|L| + 1)$  do
     $C \leftarrow C_3$ 
     $X \leftarrow \mathbf{Varlist}(C)$ 
    Repeat the main loop
end for
end function

```

The above function decomposes the original CNF formula into independent sub-formulas of variable sizes less than  $n_0$ . The set of all SAT solutions of the original formula is then the union of all solutions of all resultant sub-formulas. The set of all SAT solutions of the original formula is obtained by patching the partial assignments obtained during decomposition (gathered at the memory node) and all solutions of the resultant sub-formulas obtained at independent processing elements. The original formula is un-SAT iff all the final sub-formulas are un-SAT. This last loop can be stated as follows:

```

function GATHERALLSOLUTIONS( $C$ )
    for  $C$  at each  $PE$  do
        Compute all SAT assignments  $s(C)$  of  $C$ .
        gather  $s(C)$  at the memory node and patch the solution to the existing assignments.
    end for
end function

```

As stated above we do not discuss the actual algorithm for this function, it is assumed that such a function which computes all solutions of a small enough CNF set  $C$  is available.

#### 4.5 Computing all partial assignments

For a CNF data  $C$  in sufficiently small number of variables i.e.  $< n_0$  computation of the set of all assignments  $L$  satisfying  $C$  is required to be carried out in the algorithm. This is a well known  $\#P$  complete problem hence this loop can only be executed for a sufficiently small size of  $C$ . This problem of enumerating all assignments of a CNF formula efficiently thus requires detailed data structure and shall be discussed elsewhere.

## 4.6 Estimate of average computation time

The algorithm described above essentially only shows the decomposition of a complex CNF formula to smaller formulas with number of variables at most  $n_0$ . The final computation of all solutions is accomplished by directly solving the decomposed formulas in parallel. The computation of time complexity can thus be carried out from following estimates.

1.  $T_0$ : average time complexity of generating all solutions of CNF formula of variable size at most  $n_0$ .
2.  $d$ : maximum number of decompositions required for a formula with number of variables  $n$ , called depth,

$$\text{depth } d = N \div n_0$$

the quotient of  $N$  when divided by  $n_0$ .

3. Number of variables in the final decomposition is at most

$$r_0 = N \bmod n_0$$

4. Average time required for substitution of solutions for decomposition  $S_0$ .

Hence the average time required for solving the complete problem by decomposition is

$$T = T_0^d + d \times S_0$$

Hence the  $T_0^d$  term dominates the computation time.

## 5 Conclusions

We develop a generalization of the well known Boole-Shannon representation of Boolean functions in terms of cofactors relative to a base set of functions. For a CNF formula the base set can be chosen from the CNF factors themselves. This leads to a decomposition of the CNF formula which is shown to be useful for making a parallel solver for all satisfying assignments of the formula. However the algorithm developed on the basis of this representation is less surprising since it essentially shows substitution of partial solutions for decomposition. The intermediate step of computing all assignments of a smaller problem is a well known  $\#P$  complete problem hence  $n_0$  should necessarily be small in practice. Many heuristic improvements can be developed centered on this algorithm which will be discussed elsewhere.

### Acknowledgements

Second author greatly acknowledges support from the project grant number 11SG010 of IRCC of IIT Bombay.

## References

- [1] George Boole. An Investigation of the Laws of thought. Walton, London, 1854.
- [2] F. M. Brown. Boolean reasoning. The logic of Boolean equations. Dover, 2006.
- [3] Sergiu Rudeanu. Boolean functions and equations. North Holland, Amsterdam, 1974.
- [4] A. Biere, M. Heule, Hans van Maaren, T. Walsh (Eds). Handbook of Satisfiability. IOS Press, 2009.
- [5] Yves Crama and Peter Hammer. Boolean functions. Theory, algorithms and applications. Encyclopedia of Mathematics and its applications, vol.142. Cambridge, 2011.
- [6] Gregory Bard. Algebraic cryptanalysis. Springer 2009.
- [7] Marc Mezard and Andrea Montanari. Information, Physics and Computation. Oxford University Press, 2009.
- [8] Stuart Kauffman. Origins of order, self organization and selection in evolution. Oxford University Press, 1993.
- [9] Youssef Hammadi and C. M. Wintersteiger. Seven challenges in parallel SAT solving. Challenge paper AAAI 2012 Sub-Area spotlights track. Association of Advancement of Artificial Intelligence.
- [10] Kohavi and Jha, Switching and automata theory, Cambridge 2008.
- [11] Virendra Sule, Generalization of Boole-Shannon expansion, consistency of Boolean equations and elimination by orthonormal expansion, [arXiv.org/cs.CC/1306.2484v3](https://arxiv.org/abs/1306.2484v3), December 6, 2013.
- [12] Virendra Sule, An algorithm for Boolean satisfiability based on generalized orthonormal expansion, [arXiv.org/cs.DS/1406.4712v3](https://arxiv.org/abs/1406.4712v3), July 16, 2014.